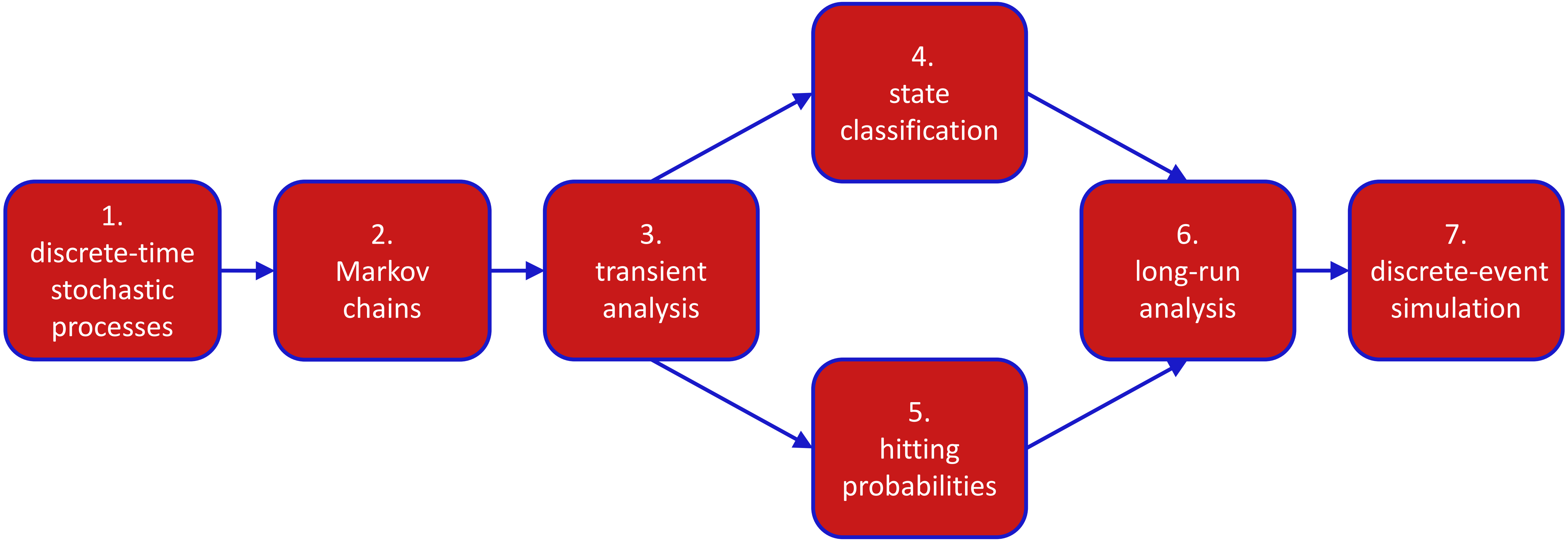
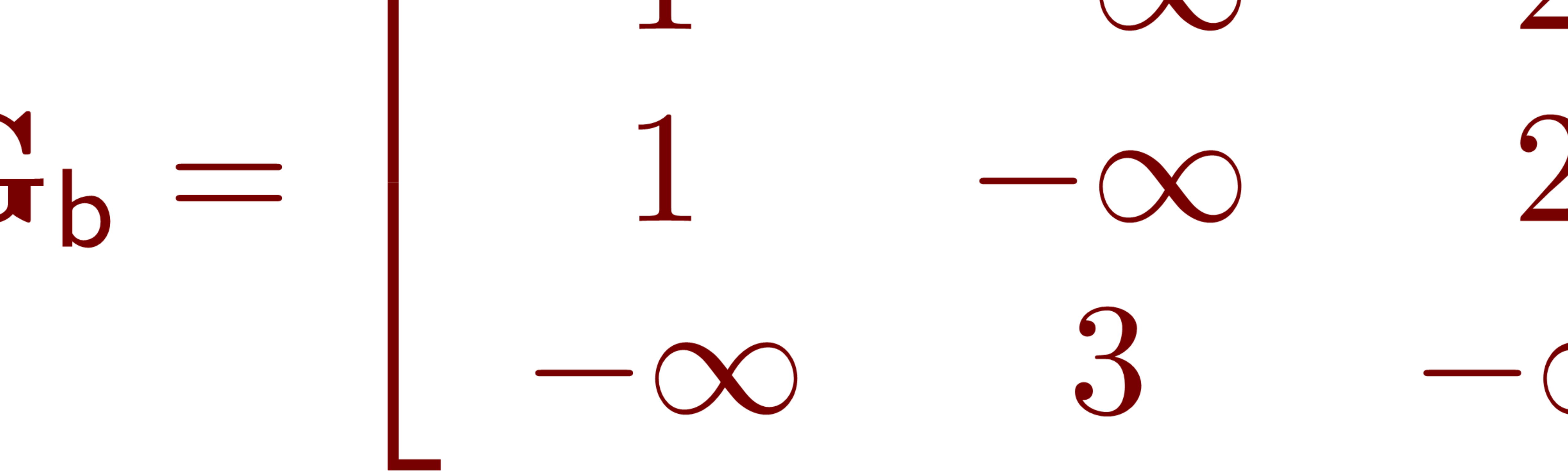
Markov modeling, discrete-event simulation – Exercises module B.7

5XIE0 Computational Modeling

Twan Basten, Marc Geilen, Jeroen Voeten
Electronic Systems Group, Department of Electrical Engineering

module B - submodules and dependencies





B.7 – discrete-event simulation

discrete-event simulation – exercises

- Section B.7 in the course notes
  - Exercise B.34 (Throughput of an Ethernetwork – confidence levels versus error bounds)
  - Exercise B.35 (Throughput of an Ethernetwork - required length of simulation sequence)
  - Exercise B.36 (Interval estimator of expected value)
  - Exercise B.37 (Confidence interval interpretation)
  - Exercise B.38 (Standard deviation - point estimator)
- answers are provided in Section B.8 of the course notes

discrete-event simulation – exercises

- section B.7 in the course notes
  - Exercise B.39 (Estimation transient distributions)
  - Exercise B.40 (Gambler's ruin – expected reward estimation)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model and compute $\pi^{(2)}$ and $E(r(X_2))$
      2. select 'Transient Reward' in 'Simulation-based Operations on Markov Chains' and enter 2 as the number of steps. As stopping criteria use a confidence level of 95% and 10,710 as the maximum number of paths. The other input fields can be left unspecified.
  - Exercise B.41 (Gambler's ruin – converge rate central limit theorem)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model (same as Exercise B.40) and for (a) compute $\pi^{(10)}$
      2. for (b) select 'Transient Distribution' in 'Simulation-based Operations on Markov Chains' and enter 10 as the number of steps. As stopping criteria use a confidence level of 95% and use 10,000 as the maximum number of paths. The other input fields are left unspecified.
      3. for (c) change the rewards assigned to the states, select 'Transient Reward' and use as stopping criteria an absolute error bound of 0.0001 and 200,000 as the maximum number of paths. The other input fields are left unspecified. To obtain more accurate estimations, leave the absolute error bound unspecified.

ES ELECTRONIC SYSTEMS TU/e

discrete-event simulation – exercises

- section B.7 in the course notes
  - Exercise B.42 (Estimation hitting probability – impact of maximal path length)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model (same as Example B.18)
      2. select 'Hitting Probability' (in 'Simulation-based Operations on Markov Chains') and select state S3. As stopping criteria use a maximal path length of 3 and use 71734 as the maximum number of paths. The other fields are left unspecified.
  - Exercise B.43 (Estimation expected cumulative reward until hit)
  - Exercise B.44 (Rover in a maze – estimation escape probability)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model (same as Exercise B.27)
      2. select 'Hitting Probability Set' and select state S29. As stopping criteria use a relative error bound of 0.01, specify the maximum number of paths to be 1000,000, and use different maximum path lengths. The other fields are left unspecified.
- answers are provided in Section B.8 of the course notes

discrete-event simulation – exercises

- section B.7 in the course notes
  - Exercise B.45 (Absolute error bound long-run expected average reward)
  - Exercise B.46 (Long-run expected average reward – confidence interval)
  - Exercise B.47 (Expected reward until return versus long-run expected average reward)
    - use CMBW (DTMC) to compute / verify answer
  - Exercise B.48 (Interval estimator long-run expected average reward)
  - Exercise B.49 (Estimation long-run expected fraction of time spent in a state)
- answers are provided in Section B.8 of the course notes

discrete-event simulation – exercises

- section B.7 in the course notes
  - Exercise B.50 (Estimation Cezàro limiting distribution of a non-unichain)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model (same as Example B.18)
      2. select ‘Cezàro Limit Distribution’ and choose ‘No preference’ regarding the recurrent state. Use the default stopping criteria.
  - Exercise B.51 (Video application – long-run average buffer occupancy)
    - use CMBW (DTMC) to compute / verify answer
      1. create the model (same as Exercise B.24) and specify the rewards
      2. select ‘Long-run Average Reward’ and choose appropriate stopping criteria
- answers are provided in Section B.8 of the course notes

Exercise B.34 (Throughput of an Ethernetwork – confidence levels versus error bounds)

Example B.20 (Throughput of an Ethernetwork). Consider a slotted Ethernetwork that behaves as a sequence of independent random variables $Y_0, Y_1, \cdots$ with a Bernoulli distribution. For each n, $P(Y_n = 1) = \mu$ and $P(Y_n = 0) = 1 - \mu$. We want to estimate parameter μ by simulation with $\gamma = 0.95$. $E(Y_n) = 1 \cdot P(Y_n = 1) + 0 \cdot P(Y_n) = 0 = \mu$, so indeed the goal is to estimate $E(Y_n)$. In a simulation the following 10 realizations are obtained: $1, 1, 0, 1, 0, 0, 1, 1, 0, 0$. The value μ is estimated by the average of the observations which is $\hat{\mu} = 0.5$. To determine a confidence interval we also need the standard deviation which is $s_{10} = 0.5$. In addition we require the value of c that corresponds to confidence level γ. $c = F^{-1}_{normal}(\frac{1+\gamma}{2}) \approx 1.96$, where F^{-1}_{normal} denotes the inverse of the distribution function of the standard normal distribution. The confidence interval $[\frac{1}{M}\sum_{n=0}^{M-1} y_n - \frac{cs_M}{\sqrt{M}}, \frac{1}{M}\sum_{n=0}^{M-1} y_n + \frac{cs_M}{\sqrt{M}}]$ is then $[0.19, 0.81]$ giving an absolute error bound of 0.31 and a relative error bound of 1.6. Such large error bounds are typically not acceptable in practice, and longer simulations are required to improve the accuracy.

$$\hat{\mu} = \frac{1}{M}\sum_{n=0}^{M-1} y_n \qquad s_M = \sqrt{\frac{1}{M}\sum_{n=0}^{M-1}\left(y_n - \frac{1}{M}\sum_{m=0}^{M-1} y_m\right)^2}$$

$$[\tfrac{1}{M}\textstyle\sum_{n=0}^{M-1} y_n - \frac{cs_M}{\sqrt{M}}, \tfrac{1}{M}\sum_{n=0}^{M-1} y_n + \frac{cs_M}{\sqrt{M}}] \tag{B.63}$$

$$|\mu - \hat{\mu}| \leq \frac{cs_M}{\sqrt{M}} \tag{B.64}$$

$$\frac{|\mu - \hat{\mu}|}{|\mu|} \leq \frac{cs_M/\sqrt{M}}{\hat{\mu} - cs_M/\sqrt{M}} \tag{B.65}$$

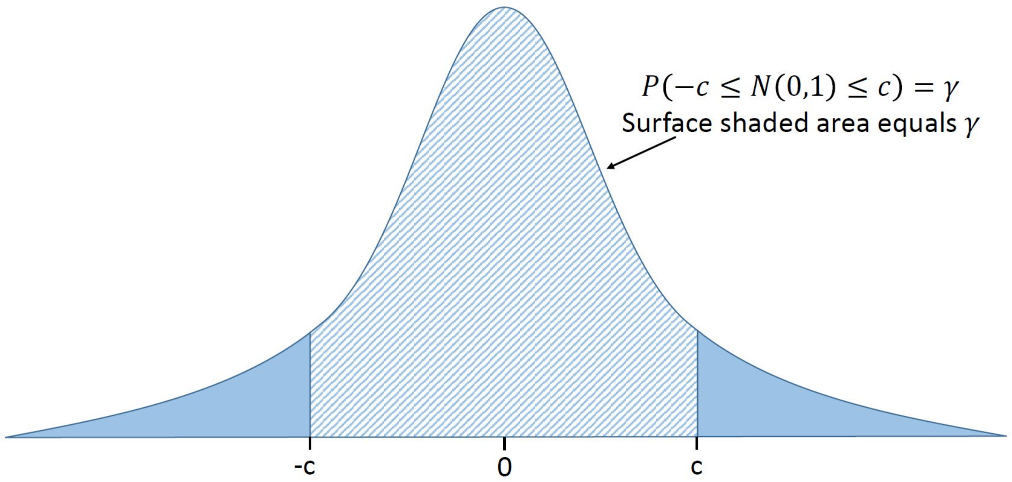


Figure B.12: Standard normal distribution

Exercise B.34 (Throughput of an Ethernetwork – confidence levels versus error bounds)

Exercise B.34 (Throughput of an Ethernetwork - confidence levels versus error bounds). Consider the Ethernet simulation of Example B.20. One can trade-off a smaller confidence level for a decrease of the estimation error bounds.

(a) Compute the confidence interval and the error bounds when a confidence level of 0.10 is taken.

(b) What is the disadvantage of taking a confidence level of 0.10?

$$|\mu - \hat{\mu}| \leq \frac{c s_M}{\sqrt{M}} \qquad \text{(B.64)}$$

$$\frac{|\mu - \hat{\mu}|}{|\mu|} \leq \frac{c s_M/\sqrt{M}}{\hat{\mu} - c s_M/\sqrt{M}} \qquad \text{(B.65)}$$

Exercise B.34 (Throughput of an Ethernetwork - confidence levels versus error bounds).

(a) For $\gamma = 0.10$, c equals $F^{-1}_{normal}((1.0 + 0.10)/2.0)$ which is approximately 0.126. The confidence interval we obtain is $[0.48, 0.51]$. The absolute and relative error bounds are 0.020 and 0.041 respectively, which are reasonable in practice.

(b) Although the estimation yields reasonable error bounds, the confidence is very low. Only 10% of confidence intervals that are obtained via simulations of length 10 contain the true value of μ. The error bounds are computed based on the assumption that μ is part of the interval, but in this case most likely it is not.

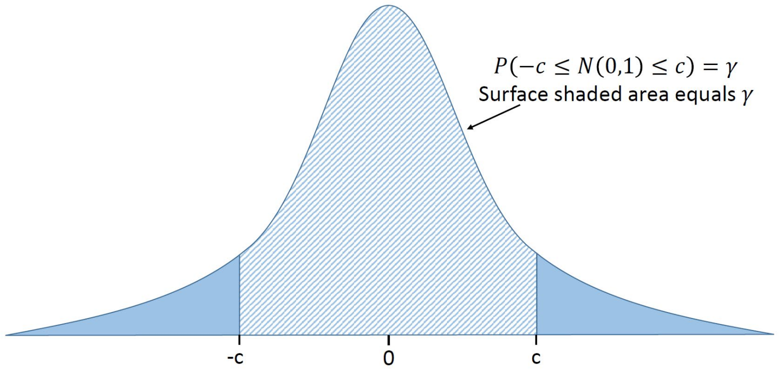


Figure B.12: Standard normal distribution

Exercise B.35 (Throughput of an Ethernetwork – required length of simulation sequence)

Exercise B.35 (Throughput of an Ethernetwork - required length of simulation sequence). Consider the Ethernet simulation of Example B.20 again. Assume the true value of μ is known to be 0.6. How long should the simulation sequence be to obtain an absolute estimation error which is at most 0.001 with a confidence level of 95%?

Exercise B.35 (Throughput of an Ethernetwork - required length of simulation sequence). Since the expected value μ is known, we can compute the variance as $\sigma^2 = Var(Y_n) = E((Y_n - \mu)^2) = E(Y_n^2 - 2\mu Y_n + \mu^2) = E(Y_n^2 - 2\mu^2 + \mu^2) = E(Y_n^2) - \mu^2$. Now $E(Y_n^2) = 1 \cdot \mu + 0 \cdot (1-\mu) = \mu$. Hence $\sigma^2 = \mu - \mu^2 = \mu(1-\mu)$ and $\sigma = \sqrt{\mu(1-\mu)}$. For $\mu = 0.6$ we thus obtain $\sigma \approx 0.775$. Now instead of using the interval estimator S_M of σ in (B.62) we can use σ itself. This leads to an absolute error bound $\frac{c\sigma}{\sqrt{M}}$. With a confidence level of 95%, $c\sigma = 1.96 * 0.775 \approx 1.52$. Hence we need $\frac{1.52}{\sqrt{M}} \leq 0.001$. This inequality holds for $M \geq 2210400$, so we require a simulation run of more than 2 million observations.

$$| \mu - \hat{\mu} | \leq \frac{c s_M}{\sqrt{M}} \tag{B.64}$$

Exercise B.36 (Interval estimator of expected value)

Exercise B.36 (Interval estimator of expected value). Derive the result of the interval estimator of μ in Equation B.62.

$$P(-c \leq \sqrt{M}\frac{\frac{1}{M}\sum_{n=0}^{M-1} Y_i - \mu}{S_M} \leq c) \approx \gamma \tag{B.61}$$

$$P(\mu \in [\frac{1}{M}\sum_{n=0}^{M-1} Y_n - \frac{cS_M}{\sqrt{M}}, \frac{1}{M}\sum_{n=0}^{M-1} Y_n + \frac{cS_M}{\sqrt{M}}]) \approx \gamma \tag{B.62}$$

Exercise B.36 (Interval estimator of expected value). We know from Equation B.61 that

$$P(-c \leq \sqrt{M}\frac{\frac{1}{M}\sum_{n=0}^{M-1} Y_i - \mu}{S_M} \leq c) \approx \gamma$$

Hence

$$P(-cS_M \leq \sqrt{M}(\frac{1}{M}\sum_{n=0}^{M-1} Y_i - \mu) \leq cS_M) \approx \gamma$$

Thus

$$P(\frac{-cS_M}{\sqrt{M}} \leq \frac{1}{M}\sum_{n=0}^{M-1} Y_i - \mu \leq \frac{cS_M}{\sqrt{M}}) \approx \gamma$$

So

$$P(\frac{-cS_M}{\sqrt{M}} - \frac{1}{M}\sum_{n=0}^{M-1} Y_i \leq -\mu \leq \frac{cS_M}{\sqrt{M}} - \frac{1}{M}\sum_{n=0}^{M-1} Y_i) \approx \gamma$$

Therefore

$$P(\frac{1}{M}\sum_{n=0}^{M-1} Y_i + \frac{cS_M}{\sqrt{M}} \geq \mu \geq \frac{1}{M}\sum_{n=0}^{M-1} Y_i - \frac{cS_M}{\sqrt{M}}) \approx \gamma$$

And thus

$$P(\frac{1}{M}\sum_{n=0}^{M-1} Y_i - \frac{cS_M}{\sqrt{M}} \leq \mu \leq \frac{1}{M}\sum_{n=0}^{M-1} Y_i + \frac{cS_M}{\sqrt{M}}) \approx \gamma$$

Rewriting yields the result:

$$P(\mu \in [\frac{1}{M}\sum_{n=0}^{M-1} Y_n - \frac{cS_M}{\sqrt{M}}, \frac{1}{M}\sum_{n=0}^{M-1} Y_n - \frac{cS_M}{\sqrt{M}}]) \approx \gamma$$

Exercise B.37 (Confidence level interpretation)

Exercise B.37 (Confidence interval interpretation). A simulation to estimate some expected value μ delivers a 95% confidence interval $[1.31, 4.25]$. A simulation expert makes the following claim: 'The probability that μ is in the interval [1.31, 4.25] is approximately 95%. Is this claim correct?

Exercise B.37 (Confidence interval interpretation). No, the claim is false. μ is either in the interval, or it is not. It can not be in the interval with probability 0.95. A correct claim would have been: 'The confidence level that μ is contained in the interval [1.31, 4.25]' is 95%'. This means that about 95% of all the intervals that could have been obtained will include value μ.

Exercise B.38 (Standard deviation – point estimator)

Exercise B.38 (Standard deviation - point estimator). Give a plausibility argument that S_M converges almost surely to σ $(=\sqrt{Var(Y_n)})$.

$$S_M = \sqrt{\frac{1}{M}\sum_{n=0}^{M-1}\left(Y_n - \frac{1}{M}\sum_{m=0}^{M-1} Y_m\right)^2}$$

Exercise B.38 (Standard deviation - point estimator). For this we have to show that for almost any sequence of realizations $y_0, y_1, \cdots$ we have

$$\lim_{M\to\infty} \sqrt{\frac{1}{M}\sum_{n=0}^{M-1}(y_n - \frac{1}{M}\sum_{m=0}^{M-1} y_m)^2} \to \sigma$$

From the strong law of large numbers we know that $\lim_{M\to\infty} \frac{1}{M}\sum_{m=0}^{M-1} y_m \to \mu$ $(= E(Y_n))$. Hence $\lim_{M\to\infty} s_M = \lim_{M\to\infty} \sqrt{\frac{1}{M}\sum_{n=0}^{M-1}(y_n - \mu)^2}$. Now $(Y_0 - \mu)^2$, $(Y_1 - \mu)^2, \cdots$ is a sequence of identically distributed random variables with mean $\sigma^2 = ((Var)(Y_n))$. Hence by the strong law of large numbers we have $\lim_{M\to\infty} \frac{1}{M}\sum_{n=0}^{M-1}(y_n - \mu)^2 = \sigma^2$. Therefore $\lim_{M\to\infty} \sqrt{\frac{1}{M}\sum_{n=0}^{M-1}(y_n - \mu)^2} = \sqrt{\sigma^2} = \sigma$.

Exercise B.39 (Estimation transient distributions)

Exercise B.39 (Estimation transient distributions). Explain how to estimate $\pi_i^{(K)}$ for some fixed $K \geq 0$ and state $i \in \mathcal{S}$.

Exercise B.39 (Estimation transient distributions). To apply the basic theory of estimation we have to define a sequence $Y_0, Y_1, \cdots$ of independent identically distributed variables such that $E(Y_n) = \pi_i^{(K)}$. Now $\pi_i^{(K)}$ is not defined as an expected value. We do know however that $E(r(X_K)) = \pi^{(K)} \cdot r^T$ (see (B.5)). If we now define r as $r(k) = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}$, then $\pi_i^{(K)} = \pi^{(K)} \cdot r^T = E(r(X_K))$. The way to estimate $E(r(X_K))$ is already explained in Example B.21.

The expected reward at time n is given by $E(r(X_n)) = \pi^{(n)} r^T$ (B.5)

Example B.21 (Estimation expected rewards). Assume we like to estimate the expected reward $E(r(X_K))$ at some time K. To apply the theory in the previous subsection, we will define a sequence of independent identically distributed variables $Y_0, Y_1, \cdots$, where each Y_i is an *independent copy* of variable $r(X_K)$. Then $E(r(X_K))$ is estimated by $\frac{1}{M}\sum_{n=0}^{M-1} y_n$ with confidence interval $[\frac{1}{M}\sum_{n=0}^{M-1} y_n - \frac{cs_M}{\sqrt{M}}, \frac{1}{M}\sum_{n=0}^{M-1} y_n - \frac{cs_M}{\sqrt{M}}]$ and corresponding bounds on absolute and relative errors. Realization y_n in the sequence $y_0, y_1, \cdots y_{M-1}$ is obtained by first (re-)starting the Markov chain by picking randomly an initial state i, based on the initial distribution (using a (pseudo)random number generator). Then a path of length K starting in state i is generated, where the state transitions are randomly chosen based on transition probabilities (also using a (pseudo)random number generator). If this paths ends in state j, $y_i = r(j)$.

Exercise B.40 (Gambler's ruin – expected reward estimation)

Exercise B.40 (Gambler's ruin - expected reward estimation). Consider the example of the gambler's ruin depicted in Figure B.3, where $\pi^{(0)} = [0, \frac{1}{2}, \frac{1}{2}, 0]$. We like to estimate $E(r(X_2))$ by simulation with an absolute estimation error bound of 2.5 and a confidence level of 95%.

(a) Estimate the number of realizations (i.e. paths) required to obtain this absolute error bound. To this end you can use the CMWB to compute $\pi^{(2)}$ and $E(r(X_2))$.

(b) Estimate $E(r(X_2))$ by simulation using the CMWB. How do you explain the difference in the obtained absolute error bound compared to (a)?

Exercise B.40 (Gambler's ruin - expected reward estimation).

(a) When the point estimation s_M is replaced by σ in Equation B.64, the absolute error bound is given by $\frac{c\sigma}{\sqrt{M}}$. For a 95% confidence interval $c \approx 1.96$. Further $\sigma^2 = Var(r(X_2)) = \{$ see also Exercise B.35 $\}$ $E(r(X_2)^2) - E(r(X_2))^2$. Now $E(r(X_2)) = 150$ and $\pi^{(2)} = [\frac{3}{8}, \frac{1}{8}, \frac{1}{8}, \frac{3}{8}]$. Hence $\sigma^2 = E(r(X_2)^2) - 150^2 = \frac{3}{8}\times 0^2 + \frac{1}{8}\times 100^2 + \frac{1}{8}\times 200^2 + \frac{3}{8}\times 300^2$ = 40,000 - 22,500 = 17,500 and thus $\sigma \approx 132$. For the absolute error bound to be 2.5 we thus need $\frac{1.96\times 132}{\sqrt{M}} = 2.5$. This holds for $M \approx 10,710$ realizations.

(b) Simulation in the CMWB could deliver estimation 149.0943 with confidence interval $[146.5805, 151.6081]$, corresponding to an absolute error bound of 2.5138. The small difference in the bound is due to the fact that the simulation uses an estimation of σ and not the true value.

$$|\mu - \hat{\mu}| \leq \frac{c s_M}{\sqrt{M}} \qquad \text{(B.64)}$$

Use CMBW (DTMC) to compute / verify answer

1. create the model and compute $\pi^{(2)}$ and $E(r(X_2))$
2. select 'Transient Reward' in 'Simulation-based Operations on Markov Chains' and enter 2 as the number of steps. As stopping criteria use a confidence level of 95% and 10,710 as the maximum number of paths. The other input fields can be left unspecified.

Exercise B.41 (Gambler's ruin – convergence rate central limit theorem)

Exercise B.41 (Gambler's ruin – converge rate central limit theorem). As a follow-up to Exercise B.40 we now like to estimate $\pi^{(10)}$ by simulation with a confidence level of 95%.

(a) Compute $\pi^{(10)}$ using the CMWB.

(b) Estimate $\pi^{(10)}$ by simulating 10,000 paths using the CMWB. Estimate the maximal absolute estimation error of the individual elements of the resulting vector, without using simulation.

(c) Estimate $\pi_2^{(10)}$ by estimating $E(r(X_{10}))$ for $r = [0, 1, 0, 0]$ using simulation with an absolute error bound of 0.0001.

Use CMBW (DTMC) to compute / verify answer

1. create the model (same as Exercise B.40) and for (a) compute $\pi^{(10)}$
2. for (b) select 'Transient Distribution' in 'Simulation-based Operations on Markov Chains' and enter 10 as the number of steps. As stopping criteria use a confidence level of 95% and use 10,000 as the maximum number of paths. The other input fields are left unspecified.
3. for (c) change the rewards assigned to the states, select 'Transient Reward' and use as stopping criteria an absolute error bound of 0.0001 and 200,000 as the maximum number of paths. The other input fields are left unspecified. To obtain more accurate estimations, leave the absolute error bound unspecified.

Exercise B.41 (Gambler's ruin – converge rate central limit theorem).

(a) $\pi^{(10)} = [0.4995, 0.0005, 0.0005, 0.4995]$.

(b) The simulation could deliver distribution $[0.4947, 0.0004, 0.0002, 0.5047]$. For each element in the vector we can estimate the absolute error bound. The first element resulted from the estimation of $E(r(X_{10}))$, where reward $r = [1, 0, 0, 0]$. The absolute error bound for this element is given by $\frac{c\sigma}{\sqrt{M}}$, where $c = 1.96$, $\sigma = \sqrt{E(r(X_{10})^2) - E(r(X_{10}))^2} = \sqrt{0.4995 - 0.4995^2} = 0.5000$ and $M = 10,000$. Hence $\frac{c\sigma}{\sqrt{M}} \approx 0.01$. The same error bound holds for the fourth element. In a similar way we can derive an error bound of about 0.00044 for the second and third elements. Hence the maximal error bound is about 0.01.

(c) When the absolute error bound of 0.0001 is used as the stopping criterion in the CMWB, the simulation will most likely terminate after 30 realization with an estimated expected reward 0.0000, confidence interval [0.0000, 0.0000] and with an absolute error bound 0.0000. How can this happen? The simulator uses the rule of thump that the distribution $\sum_{n=0}^{M-1} Y_n$ is approximately normal for $M \geq 30$ (see Equation B.60) and therefore simulates at least 30 paths. In this particular case, where each Y_n has a Bernoulli distribution with parameter 0.0005, the error between the normal distribution and binomial distribution of $\sum_{n=0}^{M-1} Y_n$ is still quite large[13]. Now point estimator $\frac{1}{M}\sum_{n=0}^{M-1} Y_n$ is used to estimate $E(r(X_{10}))$, since it converges almost surely to this value. However, the probability that $\frac{1}{30}\sum_{n=0}^{29} Y_n$ takes a value between 0.0004 and 0.0006 is 0, while the probability that it takes value 0 exceeds 0.98[14]. So after precisely 30 paths, the odds are high that only realizations with value 0 have been created, yielding the results stated above. To estimate $E(r(X_{10}))$ with an error bound of 0.0001, we could use the 'Maximum number of realizations' option in CMWB and set it to a value of about 200,000.

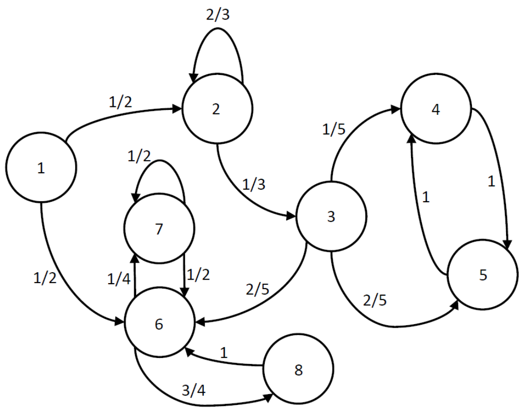
ES ELECTRONIC SYSTEMS TU/e

Exercise B.42 (Estimation hitting probability – impact of maximal path length)

Exercise B.42 (Estimation hitting probability - impact of maximal path length). Consider the transition diagram depicted in Figure B.5. We like to estimate $f_{13} = \frac{1}{2}$ by simulation, yet with a maximum path length of 3.

(a) What value will the point estimator converge to when the maximum path length is 3?

(b) Estimate the number of realizations required to obtain a relative error bound of 0.01 for a 90% confidence level. Verify you answer using the CMWB.



Use CMBW (DTMC) to compute / verify answer

1. create the model (same as Example B.18)
2. select ‘Hitting Probability’ (in ‘Simulation-based Operations on Markov Chains’) and select state S3. As stopping criteria use a maximal path length of 3 and use 71734 as the maximum number of paths. The other fields are left unspecified.

Exercise B.42 (Estimation hitting probability - impact of maximal path length).

(a) Random variable Y is defined such that Y takes value 1 if state 3 is hit from state 1 in at most 3 transitions and 0 otherwise. Two of such path that hit state 3 exist, namely $1, 2, 3$ with probability $\frac{1}{6}$ and $1, 2, 2, 3$ with probability $\frac{1}{9}$. Hence $E(Y) = \frac{5}{18}$ and thus $\frac{1}{M} \sum_{n=0}^{M-1} Y_n$ converges almost surely to $\frac{5}{18}$. Notice that the point estimation is an underestimation of f_{13}.

(b) The relative error bound is given by $\frac{c s_M/\sqrt{M}}{\hat{\mu} - c s_M/\sqrt{M}}$ (see Equation B.65). Here point estimations $\hat{\mu}$ and s_M can be replaced by $E(Y) \approx 0.28$ and $\sqrt{Var(Y)} = \sqrt{\frac{5}{18} - \frac{5}{18}^2} = \sqrt{\frac{65}{324}} \approx 0.45$ respectively. For $\gamma = 0.90$, we get $c \approx 1.65$. To obtain a relative error bound of 0.01 we thus require $\frac{1.65 \times 0.45/\sqrt{M}}{0.28 - 1.65 \times 0.45/\sqrt{M}} \leq 0.01$. This holds for $M \geq 71734$.

Exercise B.43 (Estimation expected cumulative reward until hit)

Exercise B.43 (Estimation expected cummulative reward until hit). Think of a way to estimate $\frac{f^r_{ij}}{f_{ij}}$ for some fixed states $i, j \in \mathcal{S}$.

Exercise B.43 (Estimation expected cummulative reward until hit). We have to define a sequence of independent identically distributed variables $Y_0, Y_1, \cdots$ such that $E(Y_n) = \frac{f^r_{ij}}{f_{ij}}$. Assume we would define variable Y as the cumulative reward earned if state j is hit from i in one or more steps and 0 otherwise. Then $E(Y)$ takes value f^r_{ij} and not $\frac{f^r_{ij}}{f_{ij}}$ since all paths are taken into account, including those that do not hit state j. To make sure that $E(Y) = \frac{f^r_{ij}}{f_{ij}}$, we define Y only for those paths that actually hit j[15]. We then let each Y_n be an independent copy of Y. A realization y_n is obtained by (re-)starting the Markov chain in state i, after which a path (respecting the transition probabilities) is generated. If the path hits state j, y_n equals the cumulative reward obtained along this path. In case state j is not hit, the path is simply discarded. Similar to Example B.22, this is decided when the length of the generated path exceeds some pre-defined number K.

Exercise B.44 (Rover in a maze – estimation escape probability)

Exercise B.44 (Rover in a maze - estimation escape probability). Recall Exercise B.27 of the autonomous rover. Determine the probability that the rover finds its way out of the maze, by simulation with the CMWB only. In particular pretend that the analytical answer is unknown and do not use analytical results to estimate the number of required realizations (as we did in previous exercises).

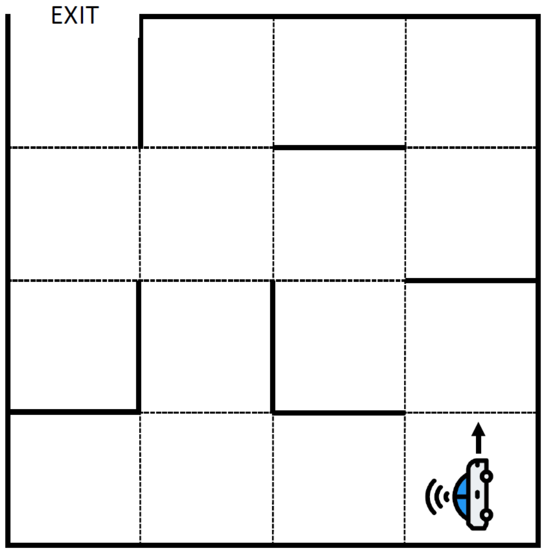


Figure B.8: An autonomous rover in a maze

Exercise B.44 (Rover in a maze - estimation escape probability). A good approach is to start experimenting with a confidence level of 95% and a relative error bound of 1%. To make sure that the simulation will terminate on the specified relative error bound, a suffiently large number of paths to generate should be specified. In this particular case 1000,000 paths will do. Since we are estimating a hitting probability, the maximal path length should be specified as well. The bigger the maximal path length, the better the estimation will be. A good way to determine an appropriate maximal path lenght is to start with a modest length, say 10, and incrementally increase the lenght by doubing it. One can continue until the estimated values are stablizing, e.g. by checking whether the most recent estimation is in the confidence interval corresponding to the prior estimation. For the rover example, the results could look like:

maximal path length	point estimation	interval estimation
10	0.1673	$[0.1657, 0.1690]$
20	0.2338	$[0.2314, 0.2361]$
40	0.2818	$[0.2790, 0.2846]$
80	0.2891	$[0.2863, 0.2920]$
160	0.2879	$[0.2851, 0.2908]$

As a resulting point estimation we thus obtain 0.2879, which is very close to the computed value of $\frac{11}{38} \approx 0.2895$.

Exercise B.45 (Absolute error bound long-run expected average reward)

Exercise B.45 (Absolute error bound long-run expected average reward). Consider the simulation technique to estimate the long-run expected average reward.

(a) Derive a bound of the absolute estimation error.

(b) Show that this bound converges to 0.

Exercise B.45 (Absolute error bound long-run expected average reward).

(a) The point estimation for μ is given by $\hat{\mu} = \frac{\sum_{n=0}^{M-1} r_n}{\sum_{n=0}^{M-1} l_n}$ and the confidence interval is given by $[\hat{\mu} - \frac{c s_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}, \hat{\mu} + \frac{c s_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}]$. Notice that $\hat{\mu}$ lies in the middle of this interval. To determine an error bound, we have to assume that μ is in the interval. It is then easy to see that $|\mu - \hat{\mu}| \leq \frac{c s_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}$.

(b) Value c is a constant and s_M converges almost surely to $\sqrt{Var(Y_n)}$. $\frac{1}{M}\sum_{n=0}^{M-1} l_n$ converges almost surely to $E(L_n) = \frac{1}{\pi_{i_r}^{(\infty)}}$. Therefore the bound converges almost surely to 0.

$$[\hat{\mu} - \frac{c s_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}, \hat{\mu} + \frac{c s_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}] \quad \text{(B.69)}$$

point estimation $\frac{\sum_{n=0}^{M-1} r_n}{\sum_{n=0}^{M-1} l_n}$ which we will denote by $\hat{\mu}$

Exercise B.46 (Long-run expected average reward – confidence interval)

Exercise B.46 (Long-run expected average reward - confidence interval). Consider a unichain with a reward function. A simulation delivers the following sequence of rewards: $1, 1, 2, 3, 3, 2, 4, 2, 4, 2, 3, 3, 3, 2, 3$. The state in which reward 2 is obtained is a recurrent state. Compute a 95% confidence interval.

Exercise B.46 (Long-run expected average reward - confidence interval). The sequence contains the following segments:

- $1, 1$: this initial segment is discarded;
- $2, 3, 3$: this is the zeroed cycle through the recurrent state; $l_0 = 3$ and $r_0 = 8$;
- $2, 4$: this is the first cycle through the recurrent state; $l_1 = 2$ and $r_1 = 6$;
- $2, 4$: this is the second cycle through the recurrent state; $l_2 = 2$ and $r_2 = 6$;
- $2, 3, 3, 3$: this is the third cycle through the recurrent state; $l_3 = 4$ and $r_3 = 11$;
- $2, 3$: this is the start of the fourth cycle; it can not be determined whether it is complete already and is therefore not (yet) taken into account.

We therefore have $M = 4$, $\hat{\mu} = \frac{\sum_{n=0}^{3} r_i}{\sum_{n=0}^{3} l_i} = \frac{31}{11} \approx 2.82$, $s_4 = \sqrt{\frac{1}{4}\sum_{n=0}^{3}(r_n - \frac{\sum_{m=0}^{3} r_m}{\sum_{m=0}^{3} l_m} l_n)^2} \approx 0.37$ and $\frac{1}{4}\sum_{n=0}^{3} l_n = 2.75$. Further for $\gamma = 0.95$, $c = 1.96$. The confidence interval is therefore $[\hat{\mu} - \frac{cs_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}, \hat{\mu} + \frac{cs_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n}] \approx [2.69, 2.95]$ with an absolute error bound $\frac{cs_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} l_n} \approx 0.13$.

Exercise B.47 (Expected reward until return versus long-run expected average reward)

Exercise B.47 (Expected reward until return versus long-run expected average reward). One can show that for any recurrent state i, $f^r_{ii} = \pi^{(\infty)} \cdot r^T / \pi^{(\infty)}_i$. This property was used as a basis to establish Equation B.66. Show that this property holds for state 2 of the Markov chain with transition probability matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{5}{6} & 0 & \frac{1}{6} \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

and reward function r defined by $r(1) = 3$, $r(2) = 5$, $r(3) = 8$ and $r(4) = 2$.

$$\frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} \text{ converges almost surely to } \mu \qquad \text{(B.66)}$$

Exercise B.47 (Expected return time versus long-run expected average reward). The chain is a non-ergodic unichain. Solving the balance equations yields $\pi^{(\infty)} = [\frac{1}{14}, \frac{3}{7}, \frac{3}{7}, \frac{1}{14}]$ and thus $\pi^{(\infty)}_2 = \frac{3}{7}$. Hence $\pi^{(\infty)} \cdot r^T = \frac{83}{14}$. Solving the equations to compute the expected reward until hitting state 2 yields $f^r_{22} = \frac{83}{6}$. The result follows from the fact that $\frac{83}{6} = \frac{83}{14} / \frac{3}{7}$.

Use CMBW (DTMC) to compute / verify answer

Exercise B.48 (Interval estimator long-run expected average reward)

Exercise B.48 (Interval estimator long-run expected average reward). Derive the interval estimator for $\pi^{(\infty)} \cdot r^T$ given in Equation B.68.

$$[\frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} - \frac{cS_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} L_n}, \frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} + \frac{cS_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} L_n}] \tag{B.68}$$

Exercise B.48 (Interval estimator long-run expected average reward). From Equation B.67 we know

$$P(-c \leq \frac{\sum_{n=0}^{M-1} Y_n}{\sqrt{M}S_M} \leq c) \approx \gamma$$

Replacing Y_n by $R_n - \mu L_n$ yields:

$$P(-c \leq \frac{\sum_{n=0}^{M-1} R_n - \mu \sum_{n=0}^{M-1} L_n}{\sqrt{M}S_M} \leq c) \approx \gamma$$

Multiplying by $\sqrt{M}S_M$ and subtracting $\sum_{n=0}^{M-1} R_n$ gives:

$$P(-c\sqrt{M}S_M - \sum_{n=0}^{M-1} R_n \leq -\mu \sum_{n=0}^{M-1} L_n \leq c\sqrt{M}S_M - \sum_{n=0}^{M-1} R_n) \approx \gamma$$

Multiplying by -1 and dividing by $\sum_{n=0}^{M-1} L_n$ yields:

$$P(\frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} - \frac{c\sqrt{M}S_M}{\sum_{n=0}^{M-1} L_n} \leq \mu \leq \frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} + \frac{c\sqrt{M}S_M}{\sum_{n=0}^{M-1} L_n}) \approx \gamma$$

The results follows by dividing by $\sqrt{M}$.

$$P(-c \leq \frac{\sum_{n=0}^{M-1} Y_n}{\sqrt{M}S_M} \leq c) \approx \gamma \tag{B.67}$$

Exercise B.49 (Estimation long-run expected fraction of time spent in a state)

Exercise B.49 (Estimation long-run expected fraction of time spent in a state). Let $P^{\propto}$ be the Cezàro limiting matrix of a unichain. Explain how to estimate $P^{\propto}_{ij}$ for fixed states $i, j \in \mathcal{S}$.

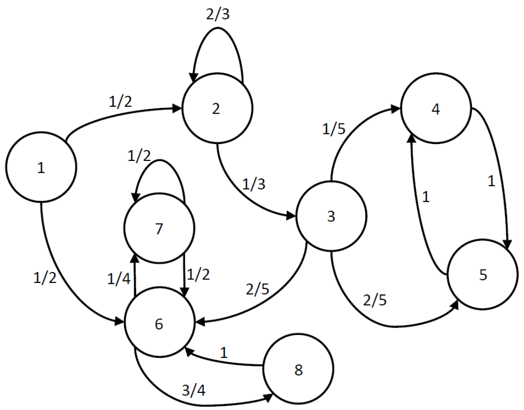
Exercise B.49 (Estimation long-run expected fraction of time spent in a state). Because the chain is a unichain, all rows in $P^{\propto}$ are equal to $\pi^{(\propto)}$. Hence $P^{\propto}_{ij} = \pi^{(\propto)}_j$. If we now define reward r as $r(k) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}$, then $\pi^{(\propto)}_j = \pi^{(\propto)} \cdot r^T$. Expressions for the point estimator and interval estimator of $\pi^{(\propto)} \cdot r^T$ are already known and given in Equations B.66 and B.68 respectively.

$$\frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} \text{ converges almost surely to } \mu \tag{B.66}$$

$$\left[\frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} - \frac{cS_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} L_n}, \frac{\sum_{n=0}^{M-1} R_n}{\sum_{n=0}^{M-1} L_n} + \frac{cS_M}{\sqrt{M}\frac{1}{M}\sum_{n=0}^{M-1} L_n}\right] \tag{B.68}$$

Exercise B.50 (Estimation Cezàro limiting distribution of a non-unichain)

Exercise B.50 (Estimation Cezàro limiting distribution of a non-unichain). Consider Example B.18 in which the Cezàro limiting distribution is computed corresponding to the transition diagram in Figure B.5 for $\pi^{(0)} = [\frac{1}{4}, \frac{3}{4}, 0, 0, 0, 0, 0, 0]$. This limit is given by $\pi^{(0)} P^{\propto} = [0, 0, 0, \frac{21}{80}, \frac{21}{80}, \frac{19}{90}, \frac{19}{180}, \frac{19}{120}]$. Make a couple of attempts to estimate this limit using CMWB. Explain your findings.



Use CMBW (DTMC) to compute / verify answer

1. create the model (same as Example B.18)
2. select ‘Cezàro Limit Distribution’ and choose ‘No preference’ regarding the recurrent state. Use the default stopping criteria.

Exercise B.50 (Estimation Cezàro limiting distribution of a non-unichain). The simulation never converges to the limit distribution. The reason is that the chain is a non-unichair with two classes of recurrent states. The simulator generates a single path through this chain. The first recurrent state it encounters on this path will be used as the state for generating cycles. In the example this is state 4, state 5 or state 6. In case state 4 or state 5 is selected, recurrent class $\{4, 5\}$ is entered and the estimation will converges to $[0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0]$. This vector corresponds to rows 4 and 5 of the Cezàro limiting matrix in Example B.18. Otherwise recurrent class $\{6, 7, 8\}$ is entered and the estimation will converge to $[0, 0, 0, 0, 0, \frac{4}{9}, \frac{2}{9}, \frac{1}{3}]$. This vector corresponds to rows 6, 7 and 8 of the limiting matrix.

$$P^{\propto} = \begin{bmatrix} 0 & 0 & 0 & \frac{3}{20} & \frac{3}{20} & \frac{14}{45} & \frac{7}{45} & \frac{7}{30} \\ 0 & 0 & 0 & \frac{3}{10} & \frac{3}{10} & \frac{8}{45} & \frac{4}{45} & \frac{2}{15} \\ 0 & 0 & 0 & \frac{3}{10} & \frac{3}{10} & \frac{8}{45} & \frac{4}{45} & \frac{2}{15} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{4}{9} & \frac{2}{9} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{4}{9} & \frac{2}{9} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{4}{9} & \frac{2}{9} & \frac{1}{3} \end{bmatrix}$$

Exercise B.51 (Video application – long-run average buffer occupancy)

Exercise B.51 (Video application - long-run average buffer occupancy). Consider the Markov chain of the video application in Exercise B.24 for $p = \frac{1}{2}$. Estimate by means of CMWB the expected long-run buffer occupancy level. Use a 95% confidence level and a relative error bound of 0.01.

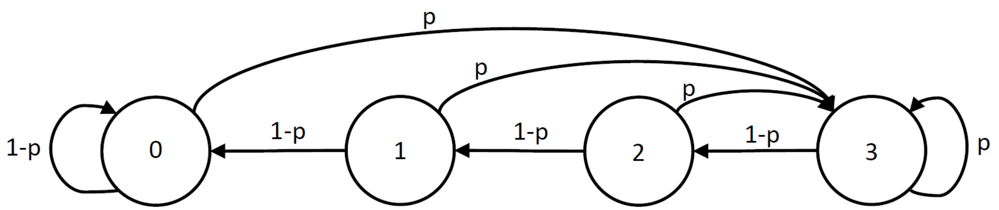


Figure B.16: Transition diagram of movie application

Exercise B.51 (Video application - long-run average buffer occupancy). To compute the occupancy level, each state i of the Markov chain can be assigned a reward i. The occupance level is then given by the long-run expected reward. The computed value equals 2.125. Simulation with a relative error bound of 0.01 could yield point estimation 2.1339 and interval estimation $[2.1127, 2.1550]$, requiring 9390 simulation cycles.

Use CMBW (DTMC) to compute / verify answer

1. create the model (same as Exercise B.24) and specify the rewards
2. select ‘Long-run Average Reward’ and choose appropriate stopping criteria